

COPY OF PAPERS
ORIGINALLY FILED



MSU 41-539.ST25

SEQUENCE LISTING

<110> Sticklen, Masomeh B
Maqbool, Shahina B
Dale, Bruce E

<120> TRANSGENIC PLANTS CONTAINING LIGNINASE AND CELLULASE WHICH
DEGRADE
LIGNIN AND CELLULOSE TO FERMENTABLE SUGARS

<130> MSU 4.1-539

<140> US/09/981,900

<141> 2001-10-18

<150> 60/242,408

<151> 2000-10-20

<160> 19

<170> PatentIn version 3.1

<210> 1

<211> 1110

<212> DNA

<213> *Oryza sativa*

<400> 1

```

gggtcggaga tgccaccacg gccacaaccc acgagcccg cgcgacacca ccgcgcgcgt      60
tgagccagcc acaaacgccc gcggataggc gcgccgcacg cggccaatcc taccacatcc      120
ccggcctccg cggctcgagc gccgtgccat ccgatccgct gagttttggc tatttatacg      180
taccgcggga gcctgtgtgc agagagtgca tctcaagaag tactcgagca aagaaggaga      240
gagcttggtg agctgcagag atggccccct ccgtgatggc gtcgtcggcc accaccgtcg      300
ctcccttcca gggctcaagt ccaccgcgg catgccgtcg cccgccgtcc gaactccagc      360
ttcggcaacg tcagcatggc ggcaggatca ggtgcatgca ggtaattacc tactgatcca      420
acacacattc ttcttcttct tcttcttctt aaccaacatt aaccaacaac tcaattatcg      480
tttattcatt gaggtgtggc cgattgaggg catcaagaag ttcgagaccc tctcctacct      540
gccaccgctc accgtggagg acctcctgaa gcagatcgag tacctagctc cgttccaagt      600
ggtgccctgc ctcgagttca gcaaggtcgg atttgtctac cgtgagaacc acaagtcccc      660
tggatactac gacggcaggt actggaccat gtggaagctg cccatgttcg ggtgcaccga      720
cgccaccag gtcgtcaagg agctcgagga ggccaagaag gcgtaccctg atgcattcgt      780
ccgtatcatc ggcttcgaca acgttaggca ggtgcagctc atcagcttca tcgcctacaa      840
cccggtctgc gaggagtctg gtggcaacta agccgtcatc gtcatatata gcctcgttta      900
attgttcac tctgattcga tgatgtctcc caccttgttt cgtgtgttcc cagtttgttt      960
catcgtcttt tgattttacc ggccgtgctc tgcttttggt ttttcttttc acctgattct     1020

```

ctctgact tgatgtaaga gtggtatctg ctacgactat atgttgtttg ggtgaggcat 1080
atgtgaatga aatctatgaa agtcccggt 1110

<210> 2
<211> 36
<212> PRT
<213> *Oryza sativa*

<400> 2
Met Ala Pro Ser Val Met Ala Ser Ser Ala Thr Thr Val Ala Pro Phe
1 5 10 15
Gln Gly Ser Ser Pro Pro Pro Ala Cys Arg Arg Pro Pro Ser Glu Leu
20 25 30
Gln Leu Arg Gln Arg Gln
35

<210> 3
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> Signal peptide targets the peroxisomes of plants

<400> 3
Arg Ala Val Ala Arg Leu
1 5

<210> 4
<211> 3004
<212> DNA
<213> *Acidothermus cellulolyticus*

<220>
<221> CDS
<222> (824)..(2512)
<223> E I beta-1,4-endoglucanase precursor

<400> 4
ggatccacgt tgtacaaggt cacctgtccg tcgttctggt agagcggcgg gatggtcacc 60
cgcacgatct ctcttttgtt gatgtcgacg gtcacgtggt tacggtttgc ctcgcccgcg 120
attttcgcg ctcgggcttgc tccggctgtc ggggttcggt tggcgtggtg tgcggagcac 180
gccgaggcga tcccaatgag ggcaagggca agagcggagc cgatggcacg tcgggtggcc 240
gatggggtag gccgatgggg cgtggcggtc ccgcccgga cagaaccgga tgcggaatag 300
gtcacggtgc gacatgttgc cgtaccgcgg acccgatga caagggtggg tgcgcgggtc 360
gcctgtgagc tgccggtggt cgtctggatc atgggaacga tcccaccatt ccccgcaatc 420
gacgcgatcg ggagcagggc ggcgcgagcc ggaccgtgtg gtcgagccgg acgattcgcc 480
cataggtgct tgcaatgcc agcgccatgt tgtcaatccg ccaaatgcag caatgcacac 540
atggacaggg attgtgactc tgagtaatga ttggattgcc ttcttgccgc ctacgcgtta 600

cgcaggtag ggcactgtat gcggtagggtt ggcgctccag ccgtgggctg gacatgcctg	660
ctgggaactc ttgacacgtc tgggtgaacg cgcaatactc ccaacaccga tgggatcggtt	720
cccataagtt tccgtctcac aacagaatcg gtgcgccctc atgatcaacg tgaaaggagt	780
acggggggaga acagacg ₃₃ gagaaaccaa cgggggattg gcg gtg ccg cgc gca	835
Val Pro Arg Ala	
1	
ttg cgg cga gtg cct ggc tgc cgg gtg atg ctg cgg gtc ggc gtc gtc	883
Leu Arg Arg Val Pro Gly Ser Arg Val Met Leu Arg Val Gly Val Val	
5 10 15 20	
gtc gcg gtg ctg gca ttg gtt gcc gca ctc gcc aac cta gcc gtg ccg	931
Val Ala Val Leu Ala Leu Val Ala Ala Leu Ala Asn Leu Ala Val Pro	
25 30 35	
cgg ccg gct cgc gcc gcg ggc ggc ggc tat tgg cac acg agc ggc cgg	979
Arg Pro Ala Arg Ala Ala Gly Gly Gly Tyr Trp His Thr Ser Gly Arg	
40 45 50	
gag atc ctg gac gcg aac aac gtg ccg gta cgg atc gcc ggc atc aac	1027
Glu Ile Leu Asp Ala Asn Asn Val Pro Val Arg Ile Ala Gly Ile Asn	
55 60 65	
tgg ttt ggg ttc gaa acc tgc aat tac gtc gtg cac ggt ctc tgg tca	1075
Trp Phe Gly Phe Glu Thr Cys Asn Tyr Val Val His Gly Leu Trp Ser	
70 75 80	
cgc gac tac cgc agc atg ctc gac cag ata aag tgc ctc ggc tac aac	1123
Arg Asp Tyr Arg Ser Met Leu Asp Gln Ile Lys Ser Leu Gly Tyr Asn	
85 90 95 100	
aca atc cgg ctg ccg tac tct gac gac att ctc aag ccg ggc acc atg	1171
Thr Ile Arg Leu Pro Tyr Ser Asp Asp Ile Leu Lys Pro Gly Thr Met	
105 110 115	
ccg aac agc atc aat ttt tac cag atg aat cag gac ctg cag ggt ctg	1219
Pro Asn Ser Ile Asn Phe Tyr Gln Met Asn Gln Asp Leu Gln Gly Leu	
120 125 130	
acg tcc ttg cag gtc atg gac aaa atc gtc gcg tac gcc ggt cag atc	1267
Thr Ser Leu Gln Val Met Asp Lys Ile Val Ala Tyr Ala Gly Gln Ile	
135 140 145	
ggc ctg cgc atc att ctt gac cgc cac cga ccg gat tgc agc ggc cag	1315
Gly Leu Arg Ile Ile Leu Asp Arg His Arg Pro Asp Cys Ser Gly Gln	
150 155 160	
tgc gcg ctg tgg tac acg agc agc gtc tgc gag gct acg tgg att tcc	1363
Ser Ala Leu Trp Tyr Thr Ser Ser Val Ser Glu Ala Thr Trp Ile Ser	
165 170 175 180	
gac ctg caa gcg ctg gcg cag cgc tac aag gga aac ccg acg gtc gtc	1411
Asp Leu Gln Ala Leu Ala Gln Arg Tyr Lys Gly Asn Pro Thr Val Val	
185 190 195	
ggc ttt gac ttg cac aac gag ccg cat gac ccg gcc tgc tgg ggc tgc	1459
Gly Phe Asp Leu His Asn Glu Pro His Asp Pro Ala Cys Trp Gly Cys	
200 205 210	
ggc gat ccg agc atc gac tgg cga ttg gcc gcc gag ccg gcc gga aac	1507

Glu	Pro	Ser	Il	Asp	Asp	Arg	Leu	Ala	Ala	Glu	Arg	Ala	Gly	Asn		
215						220					225					
gcc	gtc	ctc	tcg	gtg	aat	ccg	aac	ctg	ctc	att	ttc	gtc	gaa	ggg	gtg	1555
Ala	Val	Leu	Ser	Val	Asn	Pro	Asn	Leu	Leu	Ile	Phe	Val	Glu	Gly	Val	
230						235					240					
cag	agc	tac	aac	gac	tcc	tac	tgg	tgg	ggc	ggc	aac	ctg	caa	gga		1603
Gln	Ser	Tyr	Asn	Gly	Asp	Ser	Tyr	Trp	Trp	Gly	Gly	Asn	Leu	Gln	Gly	
245					250					255					260	
gcc	ggc	cag	tac	ccg	gtc	gtg	ctg	aac	gtg	ccg	aac	cgc	ctg	gtg	tac	1651
Ala	Gly	Gln	Tyr	Pro	Val	Val	Leu	Asn	Val	Pro	Asn	Arg	Leu	Val	Tyr	
				265					270						275	
tcg	gcg	cac	gac	tac	gcg	acg	agc	gtc	tac	ccg	cag	acg	tgg	ttc	agc	1699
Ser	Ala	His	Asp	Tyr	Ala	Thr	Ser	Val	Tyr	Pro	Gln	Thr	Trp	Phe	Ser	
			280					285						290		
gat	ccg	acc	ttc	ccc	aac	aac	atg	ccc	ggc	atc	tgg	aac	aag	aac	tgg	1747
Asp	Pro	Thr	Phe	Pro	Asn	Asn	Met	Pro	Gly	Ile	Trp	Asn	Lys	Asn	Trp	
		295					300					305				
gga	tac	ctc	ttc	aat	cag	aac	att	gca	ccg	gta	tgg	ctg	ggc	gaa	ttc	1795
Gly	Tyr	Leu	Phe	Asn	Gln	Asn	Ile	Ala	Pro	Val	Trp	Leu	Gly	Glu	Phe	
	310					315					320					
ggt	acg	aca	ctg	caa	tcc	acg	acc	gac	cag	acg	tgg	ctg	aag	acg	ctc	1843
Gly	Thr	Thr	Leu	Gln	Ser	Thr	Thr	Asp	Gln	Thr	Trp	Leu	Lys	Thr	Leu	
325					330					335					340	
gtc	cag	tac	cta	cgg	ccg	acc	gcg	caa	tac	ggg	gcg	gac	agc	ttc	cag	1891
Val	Gln	Tyr	Leu	Arg	Pro	Thr	Ala	Gln	Tyr	Gly	Ala	Asp	Ser	Phe	Gln	
				345					350					355		
tgg	acc	ttc	tgg	tcc	tgg	aac	ccc	gat	tcc	ggc	gac	aca	gga	gga	att	1939
Trp	Thr	Phe	Trp	Ser	Trp	Asn	Pro	Asp	Ser	Gly	Asp	Thr	Gly	Gly	Ile	
			360					365						370		
ctc	aag	gat	gac	tgg	cag	acg	gtc	gac	aca	gta	aaa	gac	ggc	tat	ctc	1987
Leu	Lys	Asp	Asp	Trp	Gln	Thr	Val	Asp	Thr	Val	Lys	Asp	Gly	Tyr	Leu	
		375					380					385				
gcg	ccg	atc	aag	tcg	tcg	att	ttc	gat	cct	gtc	ggc	gcg	tct	gca	tcg	2035
Ala	Pro	Ile	Lys	Ser	Ser	Ile	Phe	Asp	Pro	Val	Gly	Ala	Ser	Ala	Ser	
	390					395					400					
cct	agc	agt	caa	ccg	tcc	ccg	tcg	gtg	tcg	ccg	tct	ccg	tcg	ccg	agc	2083
Pro	Ser	Ser	Gln	Pro	Ser	Pro	Ser	Val	Ser	Pro	Ser	Pro	Ser	Pro	Ser	
405					410					415					420	
ccg	tcg	gcg	agt	cgg	acg	ccg	acg	cct	act	ccg	acg	ccg	aca	gcc	agc	2131
Pro	Ser	Ala	Ser	Arg	Thr	Pro	Thr	Pro	Thr	Pro	Thr	Pro	Thr	Ala	Ser	
				425					430					435		
ccg	acg	cca	acg	ctg	acc	cct	act	gct	acg	ccc	acg	ccc	acg	gca	agc	2179
Pro	Thr	Pro	Thr	Leu	Thr	Pro	Thr	Ala	Thr	Pro	Thr	Pro	Thr	Ala	Ser	
			440					445						450		
ccg	acg	ccg	tca	ccg	acg	gca	gcc	tcc	gga	gcc	cgc	tgc	acc	gcg	agt	2227
Pro	Thr	Pro	Ser	Pro	Thr	Ala	Ala	Ser	Gly	Ala	Arg	Cys	Thr	Ala	Ser	
		455					460					465				
tac	cag	gtc	aac	agc	gat	tgg	ggc	aat	ggc	ttc	acg	gta	acg	gtg	gcc	2275

470 Gln Val Asn Ser Asp Trp Gly Asn Gly Phe Thr Val Thr Val Ala
 475 480
 2323
 485 490 495 500
 2371
 505 510 515
 2419
 520 525 530
 2467
 535 540 545
 2512
 550 555 560
 2572
 2632
 2692
 2752
 2812
 2872
 2932
 2992
 3004

<210> 5
 <211> 562
 <212> PRT
 <213> Acidothermus cellulolyticus

<400> 5
 Val Pro Arg Ala Leu Arg Arg Val Pro Gly Ser Arg Val Met Leu Arg
 1 5 10 15
 Val Gly Val Val Val Ala Val Leu Ala Leu Val Ala Ala Leu Ala Asn
 20 25 30
 Leu Ala Val Pro Arg Pro Ala Arg Ala Ala Gly Gly Gly Tyr Trp His
 35 40 45
 Thr Ser Gly Arg Glu Ile Leu Asp Ala Asn Asn Val Pro Val Arg Ile
 50 55 60
 Ala Gly Ile Asn Trp Phe Gly Phe Glu Thr Cys Asn Tyr Val Val His
 65 70 75 80
 Gly Leu Trp Ser Arg Asp Tyr Arg Ser Met Leu Asp Gln Ile Lys Ser

85										90										95										
Leu	Gly	Tyr	Asn	Thr	Ile	Arg	Leu	Pro	Tyr	Ser	Asp	Asp	Ile	Leu	Lys															
			100						105					110																
Pro	Gly	Thr	Met	Pr	Asn	Ser	Ile	Asn	Phe	Tyr	Gln	Met	Asn	Gln	Asp															
			115				120						125																	
Leu	Gln	Gly	Leu	Thr	Ser	Leu	Gln	Val	Met	Asp	Lys	Ile	Val	Ala	Tyr															
			130				135					140																		
Ala	Gly	Gln	Ile	Gly	Leu	Arg	Ile	Ile	Leu	Asp	Arg	His	Arg	Pro	Asp															
						150				155					160															
Cys	Ser	Gly	Gln	Ser	Ala	Leu	Trp	Tyr	Thr	Ser	Ser	Val	Ser	Glu	Ala															
				165					170					175																
Thr	Trp	Ile	Ser	Asp	Leu	Gln	Ala	Leu	Ala	Gln	Arg	Tyr	Lys	Gly	Asn															
			180						185					190																
Pro	Thr	Val	Val	Gly	Phe	Asp	Leu	His	Asn	Glu	Pro	His	Asp	Pro	Ala															
			195				200					205																		
Cys	Trp	Gly	Cys	Gly	Asp	Pro	Ser	Ile	Asp	Trp	Arg	Leu	Ala	Ala	Glu															
			210			215					220																			
Arg	Ala	Gly	Asn	Ala	Val	Leu	Ser	Val	Asn	Pro	Asn	Leu	Leu	Ile	Phe															
			225			230				235				240																
Val	Glu	Gly	Val	Gln	Ser	Tyr	Asn	Gly	Asp	Ser	Tyr	Trp	Trp	Gly	Gly															
				245					250					255																
Asn	Leu	Gln	Gly	Ala	Gly	Gln	Tyr	Pro	Val	Val	Leu	Asn	Val	Pro	Asn															
			260					265					270																	
Arg	Leu	Val	Tyr	Ser	Ala	His	Asp	Tyr	Ala	Thr	Ser	Val	Tyr	Pro	Gln															
			275				280					285																		
Thr	Trp	Phe	Ser	Asp	Pro	Thr	Phe	Pro	Asn	Asn	Met	Pro	Gly	Ile	Trp															
			290				295				300																			
Asn	Lys	Asn	Trp	Gly	Tyr	Leu	Phe	Asn	Gln	Asn	Ile	Ala	Pro	Val	Trp															
					310					315				320																
Leu	Gly	Glu	Phe	Gly	Thr	Thr	Leu	Gln	Ser	Thr	Thr	Asp	Gln	Thr	Trp															
				325					330					335																
Leu	Lys	Thr	Leu	Val	Gln	Tyr	Leu	Arg	Pro	Thr	Ala	Gln	Tyr	Gly	Ala															
			340					345					350																	
Asp	Ser	Phe	Gln	Trp	Thr	Phe	Trp	Ser	Trp	Asn	Pro	Asp	Ser	Gly	Asp															
			355				360					365																		
Thr	Gly	Gly	Ile	Leu	Lys	Asp	Asp	Trp	Gln	Thr	Val	Asp	Thr	Val	Lys															
			370			375					380																			
Asp	Gly	Tyr	Leu	Ala	Pro	Ile	Lys	Ser	Ser	Ile	Phe	Asp	Pro	Val	Gly															
					390					395				400																
Ala	Ser	Ala	Ser	Pro	Ser	Ser	Gln	Pro	Ser	Pro	Ser	Val	Ser	Pro	Ser															
				405					410					415																
Pro	Ser	Pro	Ser	Pro	Ser	Ala	Ser	Arg	Thr	Pro	Thr	Pro	Thr	Pro	Thr															
				420				425					430																	

Pro Thr Ala Ser Pro Thr Pro Thr Leu Thr Pro Thr Ala Thr Pro Thr
 435 440 445
 Pro Thr Ala Ser Pro Thr Pro Ser Pro Thr Ala Ala Ser Gly Ala Arg
 450 455 460
 Cys Thr Ala Ser Tyr Gln Val Asn Ser Asp Trp Gly Asn Gly Phe Thr
 465 470 475 480
 Val Thr Val Ala Val Thr Asn Ser Gly Ser Val Ala Thr Lys Thr Trp
 485 490 495
 Thr Val Ser Trp Thr Phe Gly Gly Asn Gln Thr Ile Thr Asn Ser Trp
 500 505 510
 Asn Ala Ala Val Thr Gln Asn Gly Gln Ser Val Thr Ala Arg Asn Met
 515 520 525
 Ser Tyr Asn Asn Val Ile Gln Pro Gly Gln Asn Thr Thr Phe Gly Phe
 530 535 540
 Gln Ala Ser Tyr Thr Gly Ser Asn Ala Ala Pro Thr Val Ala Cys Ala
 545 550 555 560
 Ala Ser

<210> 6
 <211> 1467
 <212> DNA
 <213> Actinomyces naeslundii

<220>
 <221> CDS
 <222> (1)..(1467)
 <223> beta-glucosidase

<220>
 <221> misc_feature
 <222> (339)..(339)
 <223> nucleotide is uncertain

<220>
 <221> misc_feature
 <222> (947)..(947)
 <223> nucleotide is uncertain

<220>
 <221> misc_feature
 <222> (443)..(443)
 <223> nucleotide is uncertain

<400> 6
 atg acc gcc acg tcc act act tct aag agc aat ccg aac ttc ccc gac 48
 Met Thr Ala Thr Ser Thr Thr Ser Lys Ser Asn Pro Asn Phe Pro Asp
 1 5 10 15
 ggc ttc ctg tgg ggc ggg gcc acc gcc gcc aac cag atc gag ggc gct 96
 Gly Phe Leu Trp Gly Gly Ala Thr Ala Ala Asn Gln Ile Glu Gly Ala
 20 25 30
 tac aac gag gac ggc aag ggc ctg tcc gtc cag gac gtc atg cct cgg 144
 Tyr Asn Glu Asp Gly Lys Gly Leu Ser Val Gln Asp Val Met Pro Arg

40										45										
ggc	atc	atg	gcc	cac	ccc	acc	cag	gct	ccc	aca	ccg	gat	aac	ctt	caa	192				
Gly	Ile	Met	Ala	His	Pro	Thr	Gln	Ala	Pro	Thr	Pro	Asp	Asn	Leu	Gln					
50						55					60									
gct	cga	ggc	gat	ggg	cct	tct	acc	acc	gct	tac	gcc	gag	gac	atc	tcc	240				
Ala	Arg	Gly	Asp	Arg	Pro	Ser	Thr	Thr	Ala	Tyr	Ala	Glu	Asp	Ile	Ser					
65					70					75					80					
ctg	ttc	gcg	gag	atg	ggg	ttc	aag	gtc	ttc	cgc	ttc	tcc	atc	gcc	tgg	288				
Leu	Phe	Ala	Glu	Met	Gly	Phe	Lys	Val	Phe	Arg	Phe	Ser	Ile	Ala	Trp					
				85					90					95						
agc	cgc	atc	ttc	ccg	ctc	ggc	gac	gag	acc	gag	ccc	aac	gag	gaa	gga	336				
Ser	Arg	Ile		Pro	Leu	Gly	Asp	Glu	Thr	Glu	Pro	Asn	Glu	Glu	Gly					
			100					105					110							
ctn	gcc	ttc	tac	gac	cgg	gtc	ctc	gac	gag	ctc	gag	aag	cac	ggg	atc	384				
Xaa	Ala	Phe	Tyr	Asp	Arg	Val	Leu	Asp	Glu	Leu	Glu	Lys	His	Gly	Ile					
		115					120					125								
gag	cca	ctg	gtc	acc	atc	agc	cac	tac	gag	acc	ccg	ctg	cac	ctg	gcg	432				
Glu	Pro	Leu	Val	Thr	Ile	Ser	His	Tyr	Glu	Thr	Pro	Leu	His	Leu	Ala					
		130					135					140								
cgc	acc	tac	gnc	ggc	tgg	acc	gac	cgc	cgc	ctc	atc	ggc	ttc	ttc	gag	480				
Arg	Thr	Tyr	Xaa	Gly	Trp	Thr	Asp	Arg	Arg	Leu	Ile	Gly	Phe	Phe	Glu					
145					150					155					160					
cgc	tac	gcc	cgc	acc	ctg	ttc	gag	cgc	tat	ggc	aag	cgg	gtc	aag	tac	528				
Arg	Tyr	Ala	Arg	Thr	Leu	Phe	Glu	Arg	Tyr	Gly	Lys	Arg	Val	Lys	Tyr					
				165					170					175						
tgg	ctc	acc	ttc	aac	gag	atc	aac	tcc	gtg	ctc	cat	gag	ccc	ttc	cta	576				
Trp	Leu	Thr	Phe	Asn	Glu	Ile	Asn	Ser	Val	Leu	His	Glu	Pro	Phe	Leu					
			180					185					190							
tct	ggg	ggc	gtc	gcc	acg	ccc	aag	gac	agg	ccc	ccc	gag	cag	gac	ctc	624				
Ser	Gly	Gly	Val	Ala	Thr	Pro	Lys	Asp	Arg	Pro	Pro	Glu	Gln	Asp	Leu					
		195					200					205								
tac	cag	gcc	atc	caa	aac	gag	ctc	gtc	gcc	tcc	gcg	gcc	gcg	acc	agg	672				
Tyr	Gln	Ala	Ile	Gln	Asn	Glu	Leu	Val	Ala	Ser	Ala	Ala	Ala	Thr	Arg					
		210				215					220									
atc	gcc	cat	gag	acc	aac	ccc	gac	atc	cag	gtc	ggc	tgc	atg	atc	ctg	720				
Ile	Ala	His	Glu	Thr	Asn	Pro	Asp	Ile	Gln	Val	Gly	Cys	Met	Ile	Leu					
225					230					235					240					
gcc	gat	ccc	acc	tac	ccg	ctc	acc	cct	gat	ccc	cgg	gac	gtg	tgg	gcg	768				
Ala	Asp	Pro	Thr	Tyr	Pro	Leu	Thr	Pro	Asp	Pro	Arg	Asp	Val	Trp	Ala					
				245					250					255						
gcc	aag	cag	gca	gag	cgc	gcc	aac	tac	gcc	ttc	gga	gac	ctc	cac	gta	816				
Ala	Lys	Gln	Ala	Glu	Arg	Ala	Asn	Tyr	Ala	Phe	Gly	Asp	Leu	His	Val					
			260					265					270							
cgt	ggg	gag	tac	ccc	gga	tac	ctg	cgg	cgg	acc	ctg	cgg	gac	aag	ggc	864				
Arg	Gly	Glu	Tyr	Pro	Gly	Tyr	Leu	Arg	Arg	Thr	Leu	Arg	Asp	Lys	Gly					
		275					280					285								
atc	gag	ctg	gag	atc	acc	gag	gag	gac	cgc	gtg	ctg	ctg	cgg	gag	cac	912				
Ile	Glu	Leu	Glu	Ile	Thr	Glu	Glu	Asp	Arg	Val	Leu	Leu	Arg	Glu	His					

<220> The 'Xaa' at location 148 stands for Asp, Gly, Ala, or Val.

<220>

<221> misc feature

<222> (316)..(316)

<223> The 'Xaa' at location 316 stands for Tyr, Cys, Ser, or Phe.

<220>

<221> misc feature

<222> (339)..(339)

<223> nucleotide is uncertain

<220>

<221> misc feature

<222> (947)..(947)

<223> nucleotide is uncertain

<220>

<221> misc feature

<222> (443)..(443)

<223> nucleotide is uncertain

<400> 7

Met	Thr	Ala	Thr	Ser	Thr	Thr	Ser	Lys	Ser	Asn	Pro	Asn	Phe	Pro	Asp	
1				5					10					15		
Gly	Phe	Leu	Trp	Gly	Gly	Ala	Thr	Ala	Ala	Asn	Gln	Ile	Glu	Gly	Ala	
		20					25						30			
Tyr	Asn	Glu	Asp	Gly	Lys	Gly	Leu	Ser	Val	Gln	Asp	Val	Met	Pro	Arg	
	35					40						45				
Gly	Ile	Met	Ala	His	Pro	Thr	Gln	Ala	Pro	Thr	Pro	Asp	Asn	Leu	Gln	
	50					55					60					
Ala	Arg	Gly	Asp	Arg	Pro	Ser	Thr	Thr	Ala	Tyr	Ala	Glu	Asp	Ile	Ser	
65					70					75				80		
Leu	Phe	Ala	Glu	Met	Gly	Phe	Lys	Val	Phe	Arg	Phe	Ser	Ile	Ala	Trp	
			85						90					95		
Ser	Arg	Ile	Phe	Pro	Leu	Gly	Asp	Glu	Thr	Glu	Pro	Asn	Glu	Glu	Gly	
		100						105					110			
Xaa	Ala	Phe	Tyr	Asp	Arg	Val	Leu	Asp	Glu	Leu	Glu	Lys	His	Gly	Ile	
	115					120						125				
Glu	Pro	Leu	Val	Thr	Ile	Ser	His	Tyr	Glu	Thr	Pro	Leu	His	Leu	Ala	
	130					135					140					
Arg	Thr	Tyr	Xaa	Gly	Trp	Thr	Asp	Arg	Arg	Leu	Ile	Gly	Phe	Phe	Glu	
145				150						155					160	
Arg	Tyr	Ala	Arg	Thr	Leu	Phe	Glu	Arg	Tyr	Gly	Lys	Arg	Val	Lys	Tyr	
		165							170					175		
Trp	Leu	Thr	Phe	Asn	Glu	Ile	Asn	Ser	Val	Leu	His	Glu	Pro	Phe	Leu	
		180						185					190			
Ser	Gly	Gly	Val	Ala	Thr	Pro	Lys	Asp	Arg	Pro	Pro	Glu	Gln	Asp	Leu	
	195						200					205				
Tyr	Gln	Ala	Ile	Gln	Asn	Glu	Leu	Val	Ala	Ser	Ala	Ala	Ala	Thr	Arg	

215 220
 Ile Ala His Glu Thr Asn Pro Asp Ile Gln Val Gly Cys Met Ile Leu
 225 230 235 240
 Ala Asp Pro Thr Tyr Pro Leu Thr Pro Asp Pro Arg Asp Val Trp Ala
 245 250 255
 Ala Lys Gln Ala Glu Arg Ala Asn Tyr Ala Phe Gly Asp Leu His Val
 260 265 270
 Arg Gly Glu Tyr Pro Gly Tyr Leu Arg Arg Thr Leu Arg Asp Lys Gly
 275 280 285
 Ile Glu Leu Glu Ile Thr Glu Glu Asp Arg Val Leu Leu Arg Glu His
 290 295 300
 Thr Val Asp Phe Val Ser Phe Ser Tyr Tyr Met Xaa Val Cys Glu Thr
 305 310 315 320
 Val Thr Gln Ser Ala Glu Ala Gly Arg Gly Asn Leu Met Gly Gly Val
 325 330 335
 Pro Asn Pro Thr Leu Glu Ala Ser Glu Trp Gly Trp Gln Ile Asp Pro
 340 345 350
 Ala Gly Leu Arg Thr Ile Leu Asn Asp Tyr Trp Asp Arg Trp Gly Lys
 355 360 365
 Pro Leu Phe Ile Val Glu Asn Gly Leu Gly Ala Lys Asp Val Leu Val
 370 375 380
 Asp Gly Pro Asn Gly Pro Thr Val Glu Asp Asp Tyr Arg Ile Ala Tyr
 385 390 395 400
 Met Asn Asp His Leu Val Gln Val Ala Glu Ala Ile Ala Asp Gly Val
 405 410 415
 Glu Val Leu Gly Tyr Thr Ser Trp Gly Cys Ile Asp Leu Val Ser Ala
 420 425 430
 Ser Thr Ala Gln Met Ser Lys Arg Tyr Gly Phe Ile Tyr Val Asp Arg
 435 440 445
 Asp Asp Gly Gly Asn Gly Thr Leu Ala Arg Tyr Arg Lys Lys Ser Phe
 450 455 460
 Gly Trp Tyr Arg Asp Val Ile Ala Ser Asn Gly Ala Ser Leu Val Pro
 465 470 475 480
 Pro Val Gln Glu Pro Pro Arg Gly
 485

<210> 8
 <211> 3072
 <212> DNA
 <213> Streptococcus salivarius

 <220>
 <221> CDS
 <222> (392)..(2860)
 <223> 1,6-alpha-glucanhydrolase

aag cgt gaa ccg atg acg tac act atc ggt gat aac cca act caa gtt Lys Gly Glu Pro Met Thr Tyr Thr Ile Gly Asp Asn Pro Thr Gln Val 200 205 210 215	1036
tac tat ggt ccg ggc tat cca gat tgg caa aaa tac atc gca ggt gtc Tyr Tyr Asp Pro Ala Asn Pro Asp Trp Gln Lys Tyr Ile Ala Gly Val 215 220 225 230	1084
atg aaa tca gct atg gat cgt atg gga ttc gat ggt tgg caa ggt gat Met Lys Ser Ala Met Asp Arg Met Gly Phe Asp Gly Trp Gln Gly Asp 235 240 245	1132
aca att ggt gac aac cgt gtg act gat tat gag cac cgt aac agc aca Thr Ile Gly Asp Asn Arg Val Thr Asp Tyr Glu His Arg Asn Ser Thr 250 255 260	1180
gac gag gct gac tca cac atg atg tct gat tca tat gcg tca ttt att Asp Glu Ala Asp Ser His Met Met Ser Asp Ser Tyr Ala Ser Phe Ile 265 270 275	1228
aat gcc atg aag gac ctc atc ggt gaa aag tac tac atc aca atc aat Asn Ala Met Lys Asp Leu Ile Gly Glu Lys Tyr Tyr Ile Thr Ile Asn 280 285 290 295	1276
gat gtt aat ggt ggt aat gat gat aaa cta gcc aag gca cgt caa gat Asp Val Asn Gly Gly Asn Asp Asp Lys Leu Ala Lys Ala Arg Gln Asp 300 305 310	1324
gtt gtt tat aat gag ctt tgg aca aac ggt ggt tca gtt att cca gga Val Val Tyr Asn Glu Leu Trp Thr Asn Gly Gly Ser Val Ile Pro Gly 315 320 325	1372
cgt atg cag gtt gcc tat ggt gat ttg aaa gca cgt atc gat atg gta Arg Met Gln Val Ala Tyr Gly Asp Leu Lys Ala Arg Ile Asp Met Val 330 335 340	1420
cgc aat aaa act ggt aaa tca ctt atc gtt ggt gcc tac atg gaa gaa Arg Asn Lys Thr Gly Lys Ser Leu Ile Val Gly Ala Tyr Met Glu Glu 345 350 355	1468
cca ggg att gat tat act gtt cct ggc gga aaa gca act aac ggt gct Pro Gly Ile Asp Tyr Thr Val Pro Gly Gly Lys Ala Thr Asn Gly Ala 360 365 370 375	1516
ggt aaa gat gcc ctt gct ggt aaa cca ttg caa gct gat gcg act ctt Gly Lys Asp Ala Leu Ala Gly Lys Pro Leu Gln Ala Asp Ala Thr Leu 380 385 390	1564
ctc gta gat gcg aca gta gct gca gca ggt ggt tat cac atg tcc att Leu Val Asp Ala Thr Val Ala Ala Ala Gly Gly Tyr His Met Ser Ile 395 400 405	1612
gca gcc ctt gca aat gct aat gcg gcc ctt aac gtc ctt caa agt gcc Ala Ala Leu Ala Asn Ala Asn Ala Ala Leu Asn Val Leu Gln Ser Ala 410 415 420	1660
tat tac cca acg caa tac ctc agt gtg gct aaa gac act att cgt aag Tyr Tyr Pro Thr Gln Tyr Leu Ser Val Ala Lys Asp Thr Ile Arg Lys 425 430 435	1708
ctt tac aat tac caa cag ttt atc act gct tat gaa aat ctt ctc cgc Leu Tyr Asn Tyr Gln Gln Phe Ile Thr Ala Tyr Glu Asn Leu Leu Arg 440 445 450 455	1756

ggt gag ggt ggt tca aac agc act cag gct gta tct aca aag aat gct Gly Glu Gly Val Thr Asn Ser Thr Gln Ala Val Ser Thr Lys Asn Ala 460 465 470	1804
tct ggt gaa atc cct tct aaa gat gct ctt ggt gtg aca gga gat caa Ser Gly Glu Ile Leu Ser Lys Asp Ala Leu Gly Val Thr Gly Asp Gln 475 480 485	1852
gtt tgg aca ttt gct aaa tca gga aaa ggt ttc tca act gtt caa atg Val Trp Thr Phe Ala Lys Ser Gly Lys Gly Phe Ser Thr Val Gln Met 490 495 500	1900
att aat atg atg ggc atc aat gcg ggc tgg cat aat gaa gag ggt tat Ile Asn Met Met Gly Ile Asn Ala Gly Trp His Asn Glu Glu Gly Tyr 505 510 515	1948
gcg gac aat aaa aca ccg gac gca caa gaa aat ctc aca gtt cgt ctt Ala Asp Asn Lys Thr Pro Asp Ala Gln Glu Asn Leu Thr Val Arg Leu 520 525 530 535	1996
agc cta gca ggt aaa aca gcc caa gaa gca gct aaa att gct gat caa Ser Leu Ala Gly Lys Thr Ala Gln Glu Ala Lys Ile Ala Asp Gln 540 545 550	2044
gtc tat gtg acg tca ccg gat gat tgg gca act tca agc atg aag aag Val Tyr Val Thr Ser Pro Asp Asp Trp Ala Thr Ser Ser Met Lys Lys 555 560 565	2092
gca caa gca agc ctt gaa aca gat gaa aat ggt caa cca gtg ctt gtc Ala Gln Ala Ser Leu Glu Thr Asp Glu Asn Gly Gln Pro Val Leu Val 570 575 580	2140
att tca gtt cct aaa cta act ctt tgg aac atg ctt tat atc aag gaa Ile Ser Val Pro Lys Leu Thr Leu Trp Asn Met Leu Tyr Ile Lys Glu 585 590 595	2188
gac aca aca gca aca ccg gta gaa cca gtt act aac caa gct ggt aag Asp Thr Thr Ala Thr Pro Val Glu Pro Val Thr Asn Gln Ala Gly Lys 600 605 610 615	2236
aaa gta gat aat acc gta aca tct gaa gca agc tca gaa aca gct aaa Lys Val Asp Asn Thr Val Thr Ser Glu Ala Ser Ser Glu Thr Ala Lys 620 625 630	2284
tca gaa aat aca aca gta aat aaa ggt tca gag gct cca act gat acg Ser Glu Asn Thr Thr Val Asn Lys Gly Ser Glu Ala Pro Thr Asp Thr 635 640 645	2332
aaa cca tct gtt gaa gct cct aaa cta gat gaa aca act aaa cca gca Lys Pro Ser Val Glu Ala Pro Lys Leu Asp Glu Thr Thr Lys Pro Ala 650 655 660	2380
cca tca gtt gac gag tta gta aac tca gca gct gtt cca gtg gcg ata Pro Ser Val Asp Glu Leu Val Asn Ser Ala Ala Val Pro Val Ala Ile 665 670 675	2428
gct gtg tca gag acc gca cat gat aag aaa gat gac aac tca gta tct Ala Val Ser Glu Thr Ala His Asp Lys Lys Asp Asp Asn Ser Val Ser 680 685 690 695	2476
aat acg gat caa ggt aca gta gca tca gat tca atc act aca cca gct Asn Thr Asp Gln Gly Thr Val Ala Ser Asp Ser Ile Thr Thr Pro Ala 700 705 710	2524

tca gag gct gca agc aca gct gcc tca aca gtc tca tca gaa gta tca 2572
 Ser Glu Ala Ala Ser Thr Ala Ala Ser Thr Val Ser Ser Glu Val Ser
 715 720 725
 gaa agt gta aca gta tca tct gaa cca tca gaa act gaa aat agt tca 2620
 Glu Ser Val Thr Val Ser Ser Glu Pro Ser Glu Thr Glu Asn Ser Ser
 730 735 740
 gaa gca tca act tca gag tca gca act cca acg acg aca gca att tca 2668
 Glu Ala Ser Thr Ser Glu Ser Ala Thr Pro Thr Thr Thr Ala Ile Ser
 745 750 755
 gaa tca cat gca gta gtt gaa cca gtg gct tct ttg aca gaa tca gag 2716
 Glu Ser His Ala Val Val Glu Pro Val Ala Ser Leu Thr Glu Ser Glu
 760 765 770 775
 agt cag gca agc act agc ctt gtt tca gaa act aca agc aca att gtc 2764
 Ser Gln Ala Ser Thr Ser Leu Val Ser Glu Thr Thr Ser Thr Ile Val
 780 785 790
 tca gtt gct ccg tca gaa gta tca gaa agc aca tca gag gaa gtc atc 2812
 Ser Val Ala Pro Ser Glu Val Ser Glu Ser Thr Ser Glu Glu Val Ile
 795 800 805
 ctt atg gac tat cag aaa aca tca ata gtt gga ata gac tct ctg tag 2860
 Leu Met Asp Tyr Gln Lys Thr Ser Ile Val Gly Ile Asp Ser Leu
 810 815 820
 ctctctgcgt ctccagaaacc ttaccaagta cttctgaaac gattacagaa gcagcatcac 2920
 tcttttagcaa ctatgcaaga tattcagaaa cagcaagctc agaattctcac tctatggtag 2980
 cagcttcttc agaagtttct attgaaaaat tagcagtatc tatcttgaaa gatactgagg 3040
 gaggcttgta tgatgcaaca acaatcagaa at 3072

<210> 9
 <211> 822
 <212> PRT
 <213> Streptococcus salivarius

<400> 9
 Met Thr Val Asn Leu Thr Leu Gln His Ala Ser Glu Ile Ile Gly Gln
 1 5 10 15
 Asp Asn Val Asp Leu Thr Leu Ala Ala Gly Ala Ser Ala Lys Val Ser
 20 25 30
 Asn Leu Thr Val Ala Ser Glu Trp Leu Thr Asn Asn Thr Gly Tyr Leu
 35 40 45
 Val Thr Ile Ser Val Asn Asp Lys Ser Gly Asn Val Leu Ser Ser Lys
 50 55 60
 Arg Ala Gly Leu Ser Val Glu Asp Asp Trp Thr Val Phe Pro Arg Tyr
 65 70 75 80
 Gly Ile Val Ala Gly Ser Pro Thr Asp Gln Asn Ser Ile Leu Val Lys
 85 90 95
 Asn Leu Glu Ala Tyr Arg Lys Glu Leu Glu Leu Met Lys Ser Met Asn
 100 105 110

Ile Asn Ser Tyr Phe Phe Tyr Asp Ala Tyr Asn Glu Ala Thr Asp Pro
 115 120 125
 Phe Pro Glu Gly Val Asp Ser Phe Val Gln Lys Trp Asn Thr Trp Ser
 130 135 140
 His Thr Phe Val Asp Thr Lys Ala Val Lys Glu Leu Val Asp Gln Val
 145 150 155 160
 His Lys Ser Gly Ala Val Ala Met Leu Tyr Asn Met Ile Ser Ala Asp
 165 170 175
 Ser Asn Pro Lys Asn Pro Ala Leu Pro Leu Ala Ala Leu Ala Tyr Asn
 180 185 190
 Phe Tyr Asp Ser Phe Gly Lys Lys Gly Glu Pro Met Thr Tyr Thr Ile
 195 200 205
 Gly Asp Asn Pro Thr Gln Val Tyr Tyr Asp Pro Ala Asn Pro Asp Trp
 210 215 220
 Gln Lys Tyr Ile Ala Gly Val Met Lys Ser Ala Met Asp Arg Met Gly
 225 230 235 240
 Phe Asp Gly Trp Gln Gly Asp Thr Ile Gly Asp Asn Arg Val Thr Asp
 245 250 255
 Tyr Glu His Arg Asn Ser Thr Asp Glu Ala Asp Ser His Met Met Ser
 260 265 270
 Asp Ser Tyr Ala Ser Phe Ile Asn Ala Met Lys Asp Leu Ile Gly Glu
 275 280 285
 Lys Tyr Tyr Ile Thr Ile Asn Asp Val Asn Gly Gly Asn Asp Asp Lys
 290 295 300
 Leu Ala Lys Ala Arg Gln Asp Val Val Tyr Asn Glu Leu Trp Thr Asn
 305 310 315 320
 Gly Gly Ser Val Ile Pro Gly Arg Met Gln Val Ala Tyr Gly Asp Leu
 325 330 335
 Lys Ala Arg Ile Asp Met Val Arg Asn Lys Thr Gly Lys Ser Leu Ile
 340 345 350
 Val Gly Ala Tyr Met Glu Glu Pro Gly Ile Asp Tyr Thr Val Pro Gly
 355 360 365
 Gly Lys Ala Thr Asn Gly Ala Gly Lys Asp Ala Leu Ala Gly Lys Pro
 370 375 380
 Leu Gln Ala Asp Ala Thr Leu Leu Val Asp Ala Thr Val Ala Ala Ala
 385 390 395 400
 Gly Gly Tyr His Met Ser Ile Ala Ala Leu Ala Asn Ala Asn Ala Ala
 405 410 415
 Leu Asn Val Leu Gln Ser Ala Tyr Tyr Pro Thr Gln Tyr Leu Ser Val
 420 425 430
 Ala Lys Asp Thr Ile Arg Lys Leu Tyr Asn Tyr Gln Gln Phe Ile Thr
 435 440 445

Ala Tyr Glu Asn Leu Leu Arg Gly Glu Gly Val Thr Asn Ser Thr Gln
 450 455 460
 Ala Val Ser Thr Lys Asn Ala Ser Gly Glu Ile Leu Ser Lys Asp Ala
 465 470 475 480
 Leu Gly Val Thr Gly Asp Gln Val Trp Thr Phe Ala Lys Ser Gly Lys
 485 490 495
 Gly Phe Ser Thr Val Gln Met Ile Asn Met Met Gly Ile Asn Ala Gly
 500 505 510
 Trp His Asn Glu Glu Gly Tyr Ala Asp Asn Lys Thr Pro Asp Ala Gln
 515 520 525
 Glu Asn Leu Thr Val Arg Leu Ser Leu Ala Gly Lys Thr Ala Gln Glu
 530 535 540
 Ala Ala Lys Ile Ala Asp Gln Val Tyr Val Thr Ser Pro Asp Asp Trp
 545 550 555 560
 Ala Thr Ser Ser Met Lys Lys Ala Gln Ala Ser Leu Glu Thr Asp Glu
 565 570 575
 Asn Gly Gln Pro Val Leu Val Ile Ser Val Pro Lys Leu Thr Leu Trp
 580 585 590
 Asn Met Leu Tyr Ile Lys Glu Asp Thr Thr Ala Thr Pro Val Glu Pro
 595 600 605
 Val Thr Asn Gln Ala Gly Lys Lys Val Asp Asn Thr Val Thr Ser Glu
 610 615 620
 Ala Ser Ser Glu Thr Ala Lys Ser Glu Asn Thr Thr Val Asn Lys Gly
 625 630 635 640
 Ser Glu Ala Pro Thr Asp Thr Lys Pro Ser Val Glu Ala Pro Lys Leu
 645 650 655
 Asp Glu Thr Thr Lys Pro Ala Pro Ser Val Asp Glu Leu Val Asn Ser
 660 665 670
 Ala Ala Val Pro Val Ala Ile Ala Val Ser Glu Thr Ala His Asp Lys
 675 680 685
 Lys Asp Asp Asn Ser Val Ser Asn Thr Asp Gln Gly Thr Val Ala Ser
 690 695 700
 Asp Ser Ile Thr Thr Pro Ala Ser Glu Ala Ala Ser Thr Ala Ala Ser
 705 710 715 720
 Thr Val Ser Ser Glu Val Ser Glu Ser Val Thr Val Ser Ser Glu Pro
 725 730 735
 Ser Glu Thr Glu Asn Ser Ser Glu Ala Ser Thr Ser Glu Ser Ala Thr
 740 745 750
 Pro Thr Thr Thr Ala Ile Ser Glu Ser His Ala Val Val Glu Pro Val
 755 760 765
 Ala Ser Leu Thr Glu Ser Glu Ser Gln Ala Ser Thr Ser Leu Val Ser
 770 775 780
 Glu Thr Thr Ser Thr Ile Val Ser Val Ala Pro Ser Glu Val Ser Glu

800

Val Gly Ile Asp Ser Leu
820

<400>	10						
aaggtagcc	aagaacaata	gccgataaag	atagcctcat	taaacggaat	gagctagtag		60
gcaaagtcag	cgaatgtgta	tatataaagg	ttcgaggtcc	gtgcctccct	catgctctcc		120
ccatctactc	atcaactcag	atcctccagg	agacttgtac	accatctttt	gaggcacaga		180
aaccaatag	tcaaccgcgg	actggcatca	tgtatcggaa	gttggccgtc	atcacggcct		240
tcttgccac	agctcgtgct	cagtcggcct	gcactctcca	atcggagact	caccgcctc		300
tgacatggca	gaaatgctcg	tctggtggca	cttgcactca	acagacaggc	tccgtggtca		360
tcgacgcca	ctggcgctgg	actcacgcta	cgaacagcag	cacgaactgc	tacgatggca		420
acacttgagg	ctcgacccta	tgtcctgaca	acgagacctg	cgcgagaagc	tgctgtctgg		480
acggtgccgc	ctacgcgtcc	acgtacggag	ttaccacgag	cggtaacagc	ctctccattg		540
gctttgtcac	ccagtctgcg	cagaagaacg	ttggcgctcg	cctttacctt	atggcgagcg		600
acacgacct	ccaggaattc	accctgcttg	gcaacgagtt	ctctttcgat	gttgatgttt		660
cgcagctgcc	gtaagtgact	taccatgaac	ccctgacgta	tcttcttggt	ggctcccagc		720
tgactggcca	atttaagggt	cggcttgaac	ggagctctct	acttcgtgtc	catggacgcg		780
gatggtggcg	tgagcaagta	tcccaccaac	aacgctggcg	ccaagtacgg	cacggggtag		840
tgtgacagcc	agtgtccccg	cgatctgaag	ttcatcaatg	gccaggcca	cgttgagggc		900
tgggagccgt	catccaacaa	cgcaaacacg	ggcattggag	gacacggaag	ctgctgctct		960
gagatggata	tctgggaggc	caactccatc	tccgaggctc	ttacccccca	cccttgcacg		1020
actgtcggcc	aggagatctg	cgagggtgat	gggtgcggcg	gaacttactc	cgataacaga		1080
tatggcggca	cttgcgatcc	cgatggctgc	gactggaacc	cataccgcct	gggcaacacc		1140
agcttctacg	gccctggctc	aagctttacc	ctcgatacca	ccaagaaatt	gaccgttgtc		1200
accagttcg	agacgtcggg	tgccatcaac	cgatactatg	tccagaatgg	cgctcatttc		1260
cagcagccca	acgccgagct	tggtagttag	tctggcaacg	agctcaacga	tgattactgc		1320
acagctgagg	agacagaatt	cggcggatct	ctttctcaga	caagggcggc	ctgactcagt		1380
tcaagaaggc	tacctctggc	ggcatggttc	tggatcatgag	tctgtgggat	gatgtgagtt		1440
tgatggacaa	acatgcgcgt	tgacaaagag	tcaagcagct	gactgagatg	ttacagtact		1500

acgcccacat gctgtggctg gactccacct acccgacaaa cgagacctcc tccacacccg 1560
gtgccgtgcg cggaagctgc tccaccagct ccggtgtccc tgctcagggtc gaatctcagt 1620
ctcccaacgc caaggtc tttctcaaca tcaagttcgg acccattggc agcaccggca 1680
accctagcgg cggaacccc cccggcgga accgtggcac caccaccacc cgccgcccag 1740
ccactaccac tggaagctct cccggacctc cccagtctca ctacggccag tgcggcggta 1800
ttggctacag cgccccacg gtctgcgcca gcggcacaac ttgccagggtc ctgaaccctt 1860
actactctca gtgcctgtaa agctccgtgc gaaagcctga cgcaccggta gattcttgggt 1920
gagcccggtat catgacggcg gcgggagcta catggccccg ggtgatttat tttttttgta 1980
tctacttctg acccttttca aatatacggg caactcatct ttcactggag atgcggcctg 2040
cttggtattg cgatgttgct agcttggtgaa attgtggctt tcgaaaacac aaaacgattc 2100
cttagtagcc atgcatttta agataacgga atagaagaaa gaggaatta aaaaaaaaaa 2160
aaaaacaaac atcccgttca taaccgtag aatcgccgct cttcgtgtat cccagtacca 2220

<210> 11
<211> 1263
<212> DNA
<213> Phanerochaete chrysosporium

<220>
<221> CDS
<222> (34)..(1152)
<223> ckg4 ligninase precursor

<400> 11

gctacagctc accgtccggt ctacagcagca gca atg gcg ttc aag cag ctc ctc 54
Met Ala Phe Lys Gln Leu Leu
1 5

gca gcc ctc tcc gtc gcc ctg acc ctc cag gtc acc caa gct gcc ccg 102
Ala Ala Leu Ser Val Ala Leu Thr Leu Gln Val Thr Gln Ala Ala Pro
10 15 20

aac ctc gac aag cgc gtc gct tgc ccc gac ggc gtg cac acc gcc tcc 150
Asn Leu Asp Lys Arg Val Ala Cys Pro Asp Gly Val His Thr Ala Ser
25 30 35

aac gcg gcg tgc tgt gca tgg ttc ccg gtc ctc gat gat atc cag cag 198
Asn Ala Ala Cys Cys Ala Trp Phe Pro Val Leu Asp Asp Ile Gln Gln
40 45 50 55

aac ctc ttc cac ggt ggc cag tgc ggt gcc gag gcc cac gag gcc ctt 246
Asn Leu Phe His Gly Gly Gln Cys Gly Ala Glu Ala His Glu Ala Leu
60 65 70

cgt atg gtc ttc cac gac tcc atc gct atc tcg ccc aag ctt cag tcg 294
Arg Met Val Phe His Asp Ser Ile Ala Ile Ser Pro Lys Leu Gln Ser
75 80 85

cag ggc aag ttt ggc ggc ggc ggc gcg gac ggc tcg atc att acc ttc 342
Gln Gly Lys Phe Gly Gly Gly Gly Ala Asp Gly Ser Ile Ile Thr Phe

90	95	100	
tcc tgg atc gag acc acc tac cac ccg aac atc ggc ctc gac gag gtc Ser Ser Ile Glu Thr Thr Tyr His Pro Asn Ile Gly Leu Asp Glu Val 105 110 115			390
gtc gcc atc cag acc ccg ttc atc gcg aag cac ggc gtc acc cgt ggc Val Ala Ile Gln Lys Pro Phe Ile Ala Lys His Gly Val Thr Arg Gly 120 125 130 135			438
gac ttc atc gca ttc gct ggt gcc gtc ggc gtg agc aac tgc ccg ggc Asp Phe Ile Ala Phe Ala Gly Ala Val Gly Val Ser Asn Cys Pro Gly 140 145 150			486
gcg ccg cag atg cag ttc ttc ctt ggc cgc ccc gag gca acg cag gcc Ala Pro Gln Met Gln Phe Phe Leu Gly Arg Pro Glu Ala Thr Gln Ala 155 160 165			534
gcc ccc gac ggt ctc gtg ccc gag ccc ttc cac acc atc gat cag gtt Ala Pro Asp Gly Leu Val Pro Glu Pro Phe His Thr Ile Asp Gln Val 170 175 180			582
ctc gct cgc atg ctt gac gct ggt ggc ttc gac gag atc gag act gtc Leu Ala Arg Met Leu Asp Ala Gly Gly Phe Asp Glu Ile Glu Thr Val 185 190 195			630
tgg ctg ctc tct gcc cac tcc atc gcg gct gcg aac gac gtc gac ccg Trp Leu Leu Ser Ala His Ser Ile Ala Ala Asn Asp Val Asp Pro 200 205 210 215			678
acc atc tcc ggc ctg ccg ttc gac tcc act ccc ggc cag ttc gac tcc Thr Ile Ser Gly Leu Pro Phe Asp Ser Thr Pro Gly Gln Phe Asp Ser 220 225 230			726
cag ttc ttc gtc gag acc cag ctc cgc ggt acc gca ttc cct ggc aag Gln Phe Phe Val Glu Thr Gln Leu Arg Gly Thr Ala Phe Pro Gly Lys 235 240 245			774
act ggt atc cag ggc acc gtc atg tcc ccg ctc aag ggc gag atg cgt Thr Gly Ile Gln Gly Thr Val Met Ser Pro Leu Lys Gly Glu Met Arg 250 255 260			822
ctg cag acc gac cac ttg ttc gcg cgt gac tcg cgc acc gca tgc gag Leu Gln Thr Asp His Leu Phe Ala Arg Asp Ser Arg Thr Ala Cys Glu 265 270 275			870
tgg cag tcc ttc gtc aac aac cag acc aag ctg cag gag gac ttc cag Trp Gln Ser Phe Val Asn Asn Gln Thr Lys Leu Gln Glu Asp Phe Gln 280 285 290 295			918
ttc atc ttc acc gcg ctc tcg acc ctc ggc cac gac atg aac gcc atg Phe Ile Phe Thr Ala Leu Ser Thr Leu Gly His Asp Met Asn Ala Met 300 305 310			966
atc gac tgc tcc gag gtc atc ccc gcg ccc aag ccc gtc aac ttc ggc Ile Asp Cys Ser Glu Val Ile Pro Ala Pro Lys Pro Val Asn Phe Gly 315 320 325			1014
ccg tcg ttc ttc ccc gcc ggt aag acc cac gcc gac atc gag cag gcc Pro Ser Phe Phe Pro Ala Gly Lys Thr His Ala Asp Ile Glu Gln Ala 330 335 340			1062
tgc gca tcc acc ccg ttc ccg acc ctc atc acc gcc ccc ggt ccc tct Cys Ala Ser Thr Pro Phe Pro Thr Leu Ile Thr Ala Pro Gly Pro Ser			1110

345 300 355
 gcg tcc ttc ggc cgc ttc ccc ccg ccg ccg tcc ccc aac taa 1152
 Ala Ser Val Ala Arg Ile Pro Pro Pro Pro Ser Pro Asn
 360 365 370
 gctatgtcta tgctggact gctctcggtt ctacctcgtc ggtatcgtcg cacggttacc 1212
 tcgcggtttgc atcatgtata cctgctcgtg gaatatataa agtggcttat c 1263

 <210> 12
 <211> 372
 <212> PRT
 <213> *Phanerochaete chrysosporium*

 <400> 12
 Met Ala Phe Lys Gln Leu Leu Ala Ala Leu Ser Val Ala Leu Thr Leu
 1 5 10 15
 Gln Val Thr Gln Ala Ala Pro Asn Leu Asp Lys Arg Val Ala Cys Pro
 20 25 30
 Asp Gly Val His Thr Ala Ser Asn Ala Ala Cys Cys Ala Trp Phe Pro
 35 40 45
 Val Leu Asp Asp Ile Gln Gln Asn Leu Phe His Gly Gly Gln Cys Gly
 50 55 60
 Ala Glu Ala His Glu Ala Leu Arg Met Val Phe His Asp Ser Ile Ala
 65 70 75 80
 Ile Ser Pro Lys Leu Gln Ser Gln Gly Lys Phe Gly Gly Gly Gly Ala
 85 90 95
 Asp Gly Ser Ile Ile Thr Phe Ser Ser Ile Glu Thr Thr Tyr His Pro
 100 105 110
 Asn Ile Gly Leu Asp Glu Val Val Ala Ile Gln Lys Pro Phe Ile Ala
 115 120 125
 Lys His Gly Val Thr Arg Gly Asp Phe Ile Ala Phe Ala Gly Ala Val
 130 135 140
 Gly Val Ser Asn Cys Pro Gly Ala Pro Gln Met Gln Phe Phe Leu Gly
 145 150 155 160
 Arg Pro Glu Ala Thr Gln Ala Ala Pro Asp Gly Leu Val Pro Glu Pro
 165 170 175
 Phe His Thr Ile Asp Gln Val Leu Ala Arg Met Leu Asp Ala Gly Gly
 180 185 190
 Phe Asp Glu Ile Glu Thr Val Trp Leu Leu Ser Ala His Ser Ile Ala
 195 200 205
 Ala Ala Asn Asp Val Asp Pro Thr Ile Ser Gly Leu Pro Phe Asp Ser
 210 215 220
 Thr Pro Gly Gln Phe Asp Ser Gln Phe Phe Val Glu Thr Gln Leu Arg
 225 230 235 240
 Gly Thr Ala Phe Pro Gly Lys Thr Gly Ile Gln Gly Thr Val Met Ser

245	250	255
Pro Leu Lys Gly Glu Met Arg Leu Gln Thr Asp His Leu Phe Ala Arg		
260	265	270
Asp Ser Arg Thr Ala Cys Glu Trp Gln Ser Phe Val Asn Asn Gln Thr		
275	280	285
Lys Leu Gln Glu Asp Phe Gln Phe Ile Phe Thr Ala Leu Ser Thr Leu		
290	295	300
Gly His Asp Met Asn Ala Met Ile Asp Cys Ser Glu Val Ile Pro Ala		
305	310	315
Pro Lys Pro Val Asn Phe Gly Pro Ser Phe Phe Pro Ala Gly Lys Thr		
325	330	335
His Ala Asp Ile Glu Gln Ala Cys Ala Ser Thr Pro Phe Pro Thr Leu		
340	345	350
Ile Thr Ala Pro Gly Pro Ser Ala Ser Val Ala Arg Ile Pro Pro Pro		
355	360	365
Pro Ser Pro Asn		
370		

<210> 13
 <211> 1285
 <212> DNA
 <213> Phanerochaete chrysosporium

 <220>
 <221> CDS
 <222> (34)..(1149) <223> CKG5 ligninase precursor

 <400> 13

gtcagactct ccaacggttg cctttggaca gac atg gcc ttc aag aag ctc ctt	54
Met Ala Phe Lys Lys Leu Leu	
1 5	
gct gtt ctt acc gcc gct ctc tcc ctc cgc gct gcg cag ggt gcg gcc	102
Ala Val Leu Thr Ala Ala Leu Ser Leu Arg Ala Ala Gln Gly Ala Ala	
10 15 20	
gtc gag aag cgc gcg acc tgc tcg aac ggc aag gtc gtc ccc gcg gcg	150
Val Glu Lys Arg Ala Thr Cys Ser Asn Gly Lys Val Val Pro Ala Ala	
25 30 35	
tct tgc tgc acc tgg ttc aac gtt ctg tcc gat atc cag gag aac ctc	198
Ser Cys Cys Thr Trp Phe Asn Val Leu Ser Asp Ile Gln Glu Asn Leu	
40 45 50 55	
ttc aat ggc ggc cag tgt ggc gcc gag gct cat gag tcg atc cgt ctc	246
Phe Asn Gly Gly Gln Cys Gly Ala Glu Ala His Glu Ser Ile Arg Leu	
60 65 70	
gtc ttc cac gac gcc atc gct atc tct ccc gct atg gag ccg cag gcc	294
Val Phe His Asp Ala Ile Ala Ile Ser Pro Ala Met Glu Pro Gln Ala	
75 80 85	
agt tcg gtg cga ggc gcc gat ggt tct atc atg atc ttc gac gag atc	342
Ser Ser Val Arg Gly Ala Asp Gly Ser Ile Met Ile Phe Asp Glu Ile	

90	95	100	
gag acc aac ttc cat ccc aac atc ggt ctc gac gag atc gtc cgc ctg Glu Thr Asn Phe His Pro Asn Ile Gly Leu Asp Glu Ile Val Arg Leu 105 110 115			390
cag aag ccg ttc gt cag aag cac ggt gtc act ccc ggt gac ttc atc Gln Lys Phe Val Gln Lys His Gly Val Thr Pro Gly Asp Phe Ile 120 125 130 135			438
gcc ttc gct ggc gcg gtg gcg ctc agt aac tgc ccc ggt gct ccg cag Ala Phe Ala Gly Ala Val Ala Leu Ser Cys Pro Gly Ala Pro Gln 140 145 150			486
atg aac ttc ttc act ggt cgt gct ccg gca act cag cca gcc cct gac Met Asn Phe Phe Thr Gly Arg Ala Pro Ala Thr Gln Pro Ala Pro Asp 155 160 165			534
ggc ctc gtc cca gag ccc ttc cac tct gtt gac caa atc atc gac cgt Gly Leu Val Pro Glu Pro Phe His Ser Val Asp Gln Ile Ile Asp Arg 170 175 180			582
gtc ttc gat gcc ggt gaa ttc gat gag ctc gag ctc gtc tgg atg ctc Val Phe Asp Ala Gly Glu Phe Asp Glu Leu Glu Leu Val Trp Met Leu 185 190 195			630
tct gca cac tcc gtc gcg gct gcc aac gat atc gac ccg aac atc cag Ser Ala His Ser Val Ala Ala Asn Asp Ile Asp Pro Asn Ile Gln 200 205 210 215			678
ggc ttg ccc ttc gac tcg acc ccc ggt att ttc gat tcc cag ttc ttc Gly Leu Pro Phe Asp Ser Thr Pro Gly Ile Phe Asp Ser Gln Phe Phe 220 225 230			726
gtc gag act cag ctt gct ggc acc ggc ttc act ggc ggt tct aac aac Val Glu Thr Gln Leu Ala Gly Thr Gly Phe Thr Gly Gly Ser Asn Asn 235 240 245			774
cag ggc gag gtt tcc tcc ccg ctt cca ggc gag atg cgt ctc cag tct Gln Gly Glu Val Ser Ser Pro Leu Pro Gly Glu Met Arg Leu Gln Ser 250 255 260			822
gac ttc ctg atc gct cgt gac gcg cgc acc gcc tgc gag tgg cag tcg Asp Phe Leu Ile Ala Arg Asp Ala Arg Thr Ala Cys Glu Trp Gln Ser 265 270 275			870
ttc gtc aac aac cag tcc aag ctc gtc tcc gac ttc caa ttc atc ttc Phe Val Asn Asn Gln Ser Lys Leu Val Ser Asp Phe Gln Phe Ile Phe 280 285 290 295			918
ctc gcc ctc act cag ctc ggc cag gac ccg gat gcg atg acc gac tgc Leu Ala Leu Thr Gln Leu Gly Gln Asp Pro Asp Ala Met Thr Asp Cys 300 305 310			966
tct gct gtc atc ccc atc tcc aag ccc gcc ccg aac aac acc ccc gga Ser Ala Val Ile Pro Ile Ser Lys Pro Ala Pro Asn Asn Thr Pro Gly 315 320 325			1014
ttc tcc ttc ttc ccg ccc ggc atg acg atg gac gat gtc gag cag gct Phe Ser Phe Phe Pro Pro Gly Met Thr Met Asp Asp Val Glu Gln Ala 330 335 340			1062
tgc gcc gag acg ccc ttc ccg act ctc tcg act ctc cct ggc ccc gcg Cys Ala Glu Thr Pro Phe Pro Thr Leu Ser Thr Leu Pro Gly Pro Ala 1110			

345 350 355
 acc acc gtc gct cgc atc cct cct cct cct ggt gct taa gcagccatca 1159
 Thr Ser Val Ala Arg Ile Pro Pro Pro Pro Gly Ala
 360 365 370
 gacttcggg cacacac gg tattggcaac ggaaatttag aacgaagatc gtccagtgtt 1219
 ttgaagtaga aatgtgcttg tactgtgttaa acagctcttt tgacgaaata cactctgatt 1279
 tcgtcg 1285

 <210> 14
 <211> 371
 <212> PRT
 <213> *Phanerochaete chrysosporium*

 <400> 14
 Met Ala Phe Lys Lys Leu Leu Ala Val Leu Thr Ala Ala Leu Ser Leu
 1 5 10 15
 Arg Ala Ala Gln Gly Ala Ala Val Glu Lys Arg Ala Thr Cys Ser Asn
 20 25 30
 Gly Lys Val Val Pro Ala Ala Ser Cys Cys Thr Trp Phe Asn Val Leu
 35 40 45
 Ser Asp Ile Gln Glu Asn Leu Phe Asn Gly Gly Gln Cys Gly Ala Glu
 50 55 60
 Ala His Glu Ser Ile Arg Leu Val Phe His Asp Ala Ile Ala Ile Ser
 65 70 75 80
 Pro Ala Met Glu Pro Gln Ala Ser Ser Val Arg Gly Ala Asp Gly Ser
 85 90 95
 Ile Met Ile Phe Asp Glu Ile Glu Thr Asn Phe His Pro Asn Ile Gly
 100 105 110
 Leu Asp Glu Ile Val Arg Leu Gln Lys Pro Phe Val Gln Lys His Gly
 115 120 125
 Val Thr Pro Gly Asp Phe Ile Ala Phe Ala Gly Ala Val Ala Leu Ser
 130 135 140
 Asn Cys Pro Gly Ala Pro Gln Met Asn Phe Phe Thr Gly Arg Ala Pro
 145 150 155 160
 Ala Thr Gln Pro Ala Pro Asp Gly Leu Val Pro Glu Pro Phe His Ser
 165 170 175
 Val Asp Gln Ile Ile Asp Arg Val Phe Asp Ala Gly Glu Phe Asp Glu
 180 185 190
 Leu Glu Leu Val Trp Met Leu Ser Ala His Ser Val Ala Ala Ala Asn
 195 200 205
 Asp Ile Asp Pro Asn Ile Gln Gly Leu Pro Phe Asp Ser Thr Pro Gly
 210 215 220
 Ile Phe Asp Ser Gln Phe Phe Val Glu Thr Gln Leu Ala Gly Thr Gly
 225 230 235 240

Phe Thr Gly Gly Ser Asn Asn Gln Gly Glu Val Ser Ser Pro Leu Pro
 245 250 255
 Gly Glu Met Arg Leu Gln Ser Asp Phe Leu Ile Ala Arg Asp Ala Arg
 260 265 270
 Thr Ala Cys Glu Thr Gln Ser Phe Val Asn Asn Gln Ser Lys Leu Val
 275 280 285
 Ser Asp Phe Gln Phe Ile Phe Leu Ala Leu Thr Gln Leu Gly Gln Asp
 290 295 300
 Pro Asp Ala Met Thr Asp Cys Ser Ala Val Ile Pro Ile Ser Lys Pro
 305 310 315 320
 Ala Pro Asn Asn Thr Pro Gly Phe Ser Phe Phe Pro Pro Gly Met Thr
 325 330 335
 Met Asp Asp Val Glu Gln Ala Cys Ala Glu Thr Pro Phe Pro Thr Leu
 340 345 350
 Ser Thr Leu Pro Gly Pro Ala Thr Ser Val Ala Arg Ile Pro Pro Pro
 355 360 365
 Pro Gly Ala
 370

<210> 15
 <211> 360
 <212> DNA
 <213> Solanum tuberosum

<400> 15
 tgaccctaga cttgtccatc ttctggattg gccaaagttaa ttaatgtatg aaataaaagg 60
 atgcacacat agtgacatgc taatcactat aatgtgggca tcaaagttgt gtgttatgtg 120
 taataactaa ttatctgaat aagagaaaga gagatcatcc atatttctta tcctaaatga 180
 atgacagtgt ctttataatt ctttgatgaa cagatgcatt ttattaacca attccatata 240
 catataaata ttaatcatat ataattaata tcaattgggt agcaaaaccc aaatctagtc 300
 taggtgtgtt ttgctaatta tgggggatag agcaaaaaag aaactaacgt ctcaagaatc 360

<210> 16
 <211> 2521
 <212> DNA
 <213> Agrobacterium tumefaciens

<220>
 <221> CDS
 <222> (585)..(1826)
 <223> nopaline synthetase

<400> 16
 tagccgaccc agacgagcca agggatcttt ttggaatgct gctccgtcgt caggctttcc 60
 gacgtttggg tgggtgaaca gaagtcatta tcgtacggaa tgccaagcac tcccaggggg 120
 aaccctgtgg ttggcatgca catacaaatg gacgaacgga taaacctttt cagccctttt 180

aaatataccg ttattctaataaacgtcttt ttctcttagg tttaccgcgc aatataatcct	240
gtcaaacact gatagtttaa actgaaggcg ggaaacgaca atctgatcat gagcggagaa	300
ttaagggagt cacgttat, a ccccgccga tgacgcggga caagccgttt tacgtttgga	360
actgacaga cgcgaacgat tgaaggagcc actcagccgc gggtttctgg agtttaatga	420
gctaagcaca tacgtcagaa accattattg cgcgttcaaa agtcgcctaa ggtcactatc	480
agctagcaaa tatttcttgt caaaaatgct cactgacgt tccataaatt cccctcgga	540
tccaattaga gtctcatatt cactctcaat ccaaataatc tgca atg gca att acc	596
Met Ala Ile Thr	
1	
tta tcc gca act tct tta cct att tcc gcc gca gat cac cat ccg ctt	644
Leu Ser Ala Thr Ser Leu Pro Ile Ser Ala Ala Asp His His Pro Leu	
5 10 15 20	
ccc ttg acc gta ggt gtc ctc ggt tct ggt cac gcg ggg act gca tta	692
Pro Leu Thr Val Gly Val Leu Gly Ser Gly His Ala Gly Thr Ala Leu	
25 30 35	
gcg gct tgg ttc gcc tcc cgg cat gtt ccc acg gcg ctg tgg gca cca	740
Ala Ala Trp Phe Ala Ser Arg His Val Pro Thr Ala Leu Trp Ala Pro	
40 45 50	
gca gat cat cca gga tcg atc tca gca atc aag gcc aat gaa gga gtt	788
Ala Asp His Pro Gly Ser Ile Ser Ala Ile Lys Ala Asn Glu Gly Val	
55 60 65	
atc acc acc gag gga atg att aac ggt cca ttt agg gtc tca gcc tgt	836
Ile Thr Thr Glu Gly Met Ile Asn Gly Pro Phe Arg Val Ser Ala Cys	
70 75 80	
gat gac ctt gcc gca gtt att cgc tcc agc cgt gta ctg att att gta	884
Asp Asp Leu Ala Ala Val Ile Arg Ser Ser Arg Val Leu Ile Ile Val	
85 90 95 100	
acc cgt gcg gac gtt cac gac agc ttc gtc aac gaa ctc gcc aac ttc	932
Thr Arg Ala Asp Val His Asp Ser Phe Val Asn Glu Leu Ala Asn Phe	
105 110 115	
aac ggc gaa ctc gca aca aag gat att gtc gtc gtg tgc ggc cat ggc	980
Asn Gly Glu Leu Ala Thr Lys Asp Ile Val Val Val Cys Gly His Gly	
120 125 130	
ttc tcc atc aag tac gag aga cag ctg cga ttc aag cga ata ttc gag	1028
Phe Ser Ile Lys Tyr Glu Arg Gln Leu Arg Phe Lys Arg Ile Phe Glu	
135 140 145	
acg gat aat tcg ccc ata acg tct aag cta tcg gat caa aaa aaa tgt	1076
Thr Asp Asn Ser Pro Ile Thr Ser Lys Leu Ser Asp Gln Lys Lys Cys	
150 155 160	
aac gtc aac atc aag gaa atg aaa gcg tct ttc gga ctg tca tgt ttc	1124
Asn Val Asn Ile Lys Glu Met Lys Ala Ser Phe Gly Leu Ser Cys Phe	
165 170 175 180	
cca att cat cgc gat gat gct ggc gtg att gat cta ccc gaa gat acc	1172
Pro Ile His Arg Asp Asp Ala Gly Val Ile Asp Leu Pro Glu Asp Thr	
185 190 195	

aag aac atc ttt gcc cag cta ttt tcc gct aga atc atc tgc atc ccg Lys Asn Ile Phe Ala Gln Leu Phe Ser Ala Arg Ile Ile Cys Ile Pro 200 205 210	1220
ccg ttg caa gtg cta ttc ttt tcc aac tgt atc act cat gcg gtt ccg Pro Leu Gln Val Leu Phe Phe Ser Asn Cys Ile Thr His Ala Val Pro 215 220 225	1268
gca gtc atg aac atc gga aga ctc cgc gac cca gcc aat tct ctt act Ala Val Met Asn Ile Gly Arg Leu Arg Asp Pro Ala Asn Ser Leu Thr 230 235 240	1316
aaa aga gct gag aag tgg ctt ctt gaa cta gac gag cga acc cca cga Lys Arg Ala Glu Lys Trp Leu Leu Glu Leu Asp Glu Arg Thr Pro Arg 245 250 255 260	1364
gcc gag aag ggc ttt ttc ttt tat ggt gaa gga tcc aac act tac gtt Ala Glu Lys Gly Phe Phe Phe Tyr Gly Glu Gly Ser Asn Thr Tyr Val 265 270 275	1412
tgc aac gtc caa gag caa ata gac cac gaa cgc cgg aag gtt gcc gca Cys Asn Val Gln Glu Gln Ile Asp His Glu Arg Arg Lys Val Ala Ala 280 285 290	1460
gcg tgt gga ttg cgt ctc aat tct ctc ttg cag gaa tgc aat gat gaa Ala Cys Gly Leu Arg Leu Asn Ser Leu Leu Gln Glu Cys Asn Asp Glu 295 300 305	1508
tat gat act gac tat gaa act ttg agg gaa tac tgc cta gca ccg tca Tyr Asp Thr Asp Tyr Glu Thr Leu Arg Glu Tyr Cys Leu Ala Pro Ser 310 315 320	1556
cct cat aac gtg cat cat gca tgc cct gac aac atg gaa cat cgc tat Pro His Asn Val His His Ala Cys Pro Asp Asn Met Glu His Arg Tyr 325 330 335 340	1604
ttt tct gaa gaa tta tgc tgc ttg gag gat gtc gcg gca att gca gct Phe Ser Glu Glu Leu Cys Ser Leu Glu Asp Val Ala Ala Ile Ala Ala 345 350 355	1652
att gcc aac atc gaa cta ccc ctc acg cat gca ttc atc aat att att Ile Ala Asn Ile Glu Leu Pro Leu Thr His Ala Phe Ile Asn Ile Ile 360 365 370	1700
cat gcg ggg aaa ggc aag att aat cca act ggc aaa tca tcc agc gtg His Ala Gly Lys Gly Lys Ile Asn Pro Thr Gly Lys Ser Ser Ser Val 375 380 385	1748
att ggt aac ttc agt tcc agc gac ttg att cgt ttt ggt gct acc cac Ile Gly Asn Phe Ser Ser Ser Asp Leu Ile Arg Phe Gly Ala Thr His 390 395 400	1796
gtt ttc aat aag gac gag atg gtg gag taa agaaggagtgc cgtcgaagca Val Phe Asn Lys Asp Glu Met Val Glu 405 410	1846
gatcggttcaa acatttggca ataaagtttc ttaagattga atcctgttgc cgggtcttgcg	1906
atgattatca tataatttct gttgaattac gttaagcatg taataattaa catgtaatgc	1966
atgacgttat ttatgagatg ggtttttatg attagagtcc cgcaattata catttaatac	2026
gcgatagaaa acaaaatata gcgcgcaaac taggataaat tatcgcgcgcg ggtgtcatct	2086

atgttactag atcgatcaaa ctccggtact gtgtaatgac gatgagcaat cgagaggctg 2146
 actaacaara ggatgccc aaacaacct ctccaaactg tttcgaattg gaagtttctg 2206
 ctcattgccg caggcatc ttagatattc gcgggctatt cccactaatt cgtcctgctg 2266
 gtttgcgccc agataaa ca gtgcatctcc ttacaagttc ctctgtcttg tgaatgaac 2326
 tgctgactgc ccccaagaa agcctcctca tctcccagtt ggcggcggtg gatacaccat 2386
 cgaaaaccca cgtccgaaca cttgatacat gtgcctgaga aataggccta cgtccaagag 2446
 caagtctttt ctgtgctcgt cggaattcc tctcctgtca gacggtcgtg cgcattgttt 2506
 gcgttgatga agctt 2521

<210> 17
 <211> 413
 <212> PRT
 <213> Agrobacterium tumefaciens

<400> 17
 Met Ala Ile Thr Leu Ser Ala Thr Ser Leu Pro Ile Ser Ala Ala Asp
 1 5 10 15
 His His Pro Leu Pro Leu Thr Val Gly Val Leu Gly Ser Gly His Ala
 20 25 30
 Gly Thr Ala Leu Ala Ala Trp Phe Ala Ser Arg His Val Pro Thr Ala
 35 40 45
 Leu Trp Ala Pro Ala Asp His Pro Gly Ser Ile Ser Ala Ile Lys Ala
 50 55 60
 Asn Glu Gly Val Ile Thr Thr Glu Gly Met Ile Asn Gly Pro Phe Arg
 65 70 75 80
 Val Ser Ala Cys Asp Asp Leu Ala Ala Val Ile Arg Ser Ser Arg Val
 85 90 95
 Leu Ile Ile Val Thr Arg Ala Asp Val His Asp Ser Phe Val Asn Glu
 100 105 110
 Leu Ala Asn Phe Asn Gly Glu Leu Ala Thr Lys Asp Ile Val Val Val
 115 120 125
 Cys Gly His Gly Phe Ser Ile Lys Tyr Glu Arg Gln Leu Arg Phe Lys
 130 135 140
 Arg Ile Phe Glu Thr Asp Asn Ser Pro Ile Thr Ser Lys Leu Ser Asp
 145 150 155 160
 Gln Lys Lys Cys Asn Val Asn Ile Lys Glu Met Lys Ala Ser Phe Gly
 165 170 175
 Leu Ser Cys Phe Pro Ile His Arg Asp Asp Ala Gly Val Ile Asp Leu
 180 185 190
 Pro Glu Asp Thr Lys Asn Ile Phe Ala Gln Leu Phe Ser Ala Arg Ile
 195 200 205
 Ile Cys Ile Pro Pro Leu Gln Val Leu Phe Phe Ser Asn Cys Ile Thr

210	213	220
His Ala Val Pro Ala Val Met Asn Ile Gly Arg Leu Arg Asp Pro Ala 225 230 235 240		
Asn Ser Leu Thr Lys Arg Ala Glu Lys Trp Leu Leu Glu Leu Asp Glu 245 250 255		
Arg Thr Pro Arg Ala Glu Lys Gly Phe Phe Phe Tyr Gly Glu Gly Ser 260 265 270		
Asn Thr Tyr Val Cys Asn Val Gln Glu Gln Ile Asp His Glu Arg Arg 275 280 285		
Lys Val Ala Ala Ala Cys Gly Leu Arg Leu Asn Ser Leu Leu Gln Glu 290 295 300		
Cys Asn Asp Glu Tyr Asp Thr Asp Tyr Glu Thr Leu Arg Glu Tyr Cys 305 310 315 320		
Leu Ala Pro Ser Pro His Asn Val His His Ala Cys Pro Asp Asn Met 325 330 335		
Glu His Arg Tyr Phe Ser Glu Glu Leu Cys Ser Leu Glu Asp Val Ala 340 345 350		
Ala Ile Ala Ala Ile Ala Asn Ile Glu Leu Pro Leu Thr His Ala Phe 355 360 365		
Ile Asn Ile Ile His Ala Gly Lys Gly Lys Ile Asn Pro Thr Gly Lys 370 375 380		
Ser Ser Ser Val Ile Gly Asn Phe Ser Ser Ser Asp Leu Ile Arg Phe 385 390 395 400		
Gly Ala Thr His Val Phe Asn Lys Asp Glu Met Val Glu 405 410		

<210> 18
 <211> 835
 <212> DNA
 <213> Streptomyces hygroscopicus

<400> 18
 gctcgctgtc attttcgaga cgccatcttt ggaagcgggtg gccgaatccg tactgcgcgg 60
 actcgacgac gcgtaaaacg atcgaccacg tacacgagtc cggacacggg gcgaggaggc 120
 ccggttccgg caccgaggaa gaccgaagga agaccacacg tgagcccaga acgacgcccg 180
 gccgacatcc gccgtgccac cgaggcggac atgccggcgg tctgcacat cgtaaccac 240
 tacatcgaga caagcacggt caacttccgt accgagccgc aggaaccgca ggagtggacg 300
 gacgacctcg tccgtctgcg ggagcgtat ccctggctcg tcgccgaggt ggacggcgag 360
 gtcgccggca tcgcctacgc gggcccctgg aaggcacgca acgcctacga ctggacggcc 420
 gactcgaccg tgtacgtctc cccccccac cagcggacgg gactgggctc cagctctac 480
 acccaactgc tgaagtcctt ggaggcacag ggcttcaaga gcgtggctgc tgtcatcggt 540
 ctgccaacg acccgagcgt gcgcatgcac gaggcgtctg gatatgcccc ccgcggcatg 600

ctgcggggcgg ccggcttcaa gcacgggaac tggcatgacg tgggtttctg gcagctggac 660
 ttcagcctgc cggtagcgcc ccgtccggtc ctgcccgta ccgagatctg aacggagtgc 720
 gcgtggggcgc cggccgagct ggagctggta cgggaactca tcgaactcaa ctggcatacc 780
 cgcaatggtg aggtggaaac gggcgggatc gcgtacgacc gtgccagga ggcct 835

<210> 19
 <211> 623
 <212> DNA
 <213> *Oryza sativa*

<400> 19
 tatatacata cccccccctc tctctccatc cccccaaccc taccaccacc accaccacca 60
 cctctcccc cctcgtgccc ggacgacgag ctctctcccc ctccccctcc gccgcgcgcg 120
 gtaaccaccc cgcgtccctc tctcttttct ttctccgttt tttttttccg tctcgtctcg 180
 atctttggcc ttggtagttt gggggcgaga ggcggcttcg tcgccagat cggtgcgcg 240
 gagggggcggg atctcggggc tgggtctcgg cgtgcggccg gatcctcgcg ggaatgggg 300
 ctctcggatg tagatctgat ccgccgttgt tgggggagat gatggggcgt taaaatttc 360
 gccatgctaa acaagatcag gaagagggga aaagggcact atggtttata tttttatata 420
 tttctgctgc tgctcgtcag gcttagatgt gctagatctt tctttcttct tttgtgggt 480
 agaatttgaa tccctcagca ttgttcacg gtagtttttc tttcatgat ttgtgacaaa 540
 tgcagcctcg tgcggagctt tttgtaggt agaagatggc tgacgccgag gatgggggat 600
 ccccggtggtg tcagtcctt atg 623

<210> 20
 <211> 1506
 <212> DNA
 <213> *Trichoderma longibrachiatum*

<400> 20
 accatggctc aatctgcttg tactcttcaa tctgagactc atctccact tacttggcag 60
 aagtgttcat ctggtggtac ttgtactcaa cagactggat ctgttggtat tgatgctaac 120
 tggagatgga ctcatgctac taactcttct actaactgct atgatggtaa cacttggcca 180
 tctactcttt gtcctgataa cgagacttgt gctaagaact gctgtcttga tgggtcgtct 240
 tacgcttcta cttacggagt tactacttct ggaaactctc tttctattgg attcgttact 300
 cagtctgctc agaagaacgt tgggtgctagg ttgtacttga tggcttctga tactacttac 360
 caagattca ctcttcttg taacgagttc tctttcgatg ttgatgttc tcaacttcca 420
 tgtggtttga acggagcttt gtacttcgtt tctatggatg ctgatgggtg agtttctaag 480
 tatccaacta aactgctgg agctaagtat ggtactgggt actgtgatcc tcagtgtcca 540
 agagatctta agttcattaa cggacaagct aatgttgagg gatgggagcc atcttctaac 600

aatgctaaca ctggtattgg aggcacatgga tcttgttgct ctgagatgga tatttgggag 660
gctaacteta tctctgagcc tcttactcca catccatgca ctactgttgg acaagagatt 720
tgcgagggag atggttgttg tggaacttac tctgataaca gatacggagg tacttgcgat 780
ccagatggag gtgattgga tccatacaga cttggtaaca cttctttcta cggccagga 840
tcttcattca ctcttgatac tactaagaag ttgactgttg ttactcagtt cgagacttct 900
ggtgctatca acagatacta cgttcagaat ggagttactt tccaacaacc taacgctgag 960
cttggttctt actctggtaa cgagttgaac gatgattact gtactgctga ggaagctgag 1020
tttggtggat catctttctc tgataagggt ggacttactc agttcaagaa ggctacttct 1080
ggtggtatgg ttcttgttat gtctcttgg gatgattact acgctaacat gttgtggctt 1140
gattctactt acccaactaa cgagacttct tctactccag gtgctgttag aggatcttgc 1200
tctacttctt ctggtgttcc tgctcaagtt gaatctcaat ctccaatgc taaggttact 1260
ttctctaaca tcaagttcgg accaattgga tctactggta acccttctgg aggtaatcca 1320
cctggaggta atccacctgg aactactaca actaggagac cagctactac aactggatca 1380
tctccaggac ctactcaatc tcattacggt caatgtggag gtattgggta ctctgggtcca 1440
actgtttgtg cttctggaac tacttgtcaa gttcttaacc cttactattc tcaatgcctt 1500
taatga 1506

<210> 21
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR primer SP1F

<400> 21
ccgcctaggc gcatggcccc ctccgt 26

<210> 22
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR primer SP3R

<400> 22
cgctgtacac gcacctgata ctgcc 25